

APPENDIX A: WEIGHTING, VARIANCE ESTIMATION, & STATISTICAL TESTING

Weighting & Variance Estimation

School, classroom and students data were weighted to produce total population estimates. The weighting factor reflects the probability of selection, non-response, and post-stratification (gender & form). Variances were estimated using the general linear variance estimators. This method of computing variances takes into account the complex nature of the design and the classroom effect. It also accounts for sampling with the probability proportional to measure size. SUDAAN was used to compute standard errors for the estimates

A percent and its estimated standard error may be used to construct confidence intervals (C.I.) about the percent. The C.I. is expressed as a range (upper and lower) around the percent. The C.I. range contains the average value of the percent, which would result if all possible samples were produced. The 95% C.I. suggests that if 100 samples were drawn and C.I.s were calculated for each, then the average value of the percent would be contained in 95 of the 100 C.I.s

Statistical Testing

The test of statistical significance is done by comparing the 95% C.I. for two percentages. If the C.I.s do not overlap then the percentages are significantly different.

E.g. In the table on prevalence, 30.1% males and 21.5% females from Harare had ever tried smoking.

- The 95% C.I. for each percent is calculated by multiplying the standard error (SE) by 1.96, giving 6.9 & 4.0 for males and females respectively.
- Therefore the lower and upper bounds for the two percentages are
Males 30.1% = [23.2,37.0]
&
Females 21.5% = [17.5,25.5]
- Statistical difference is determined by comparing the upper bound, for the smaller % and the lower bound, for the larger %
 - if the two ranges do not overlap, then the two %s are statistically (significantly) different at the 95% C.I.
 - if the two ranges overlap, then there is no statistical (significant) difference between the two %s, at the 95% C.I.
- In this example, the percentages 30.1 [23.2, 37.0] and 21.5 [17.5, 25.5] overlap, therefore there is no significant difference in students who have ever smoked, between male and female students.